

Draft Genome Sequence of *Pantoea ananatis* Strain LMG 2665^T, a Bacterial Pathogen of Pineapple Fruitlets

Zaky Adam, James T. Tambong, Christopher T. Lewis, C. André Lévesque, Wen Chen, Eden S. P. Bromfield, Izhar U. H. Khan, Renlin Xu

Eastern Cereal and Oilseed Research Center (ECORC), Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada

We report the draft genome sequence of *Pantoea ananatis* LMG 2665^T, the bacterial causal agent of pineapple fruitlet rot.

Received 2 May 2014 Accepted 6 May 2014 Published 22 May 2014

Citation Adam Z, Tambong JT, Lewis CT, Lévesque CA, Chen W, Bromfield ESP, Khan IUH, Xu R. 2014. Draft genome sequence of *Pantoea ananatis* strain LMG 2665^T, a bacterial pathogen of pineapple fruitlets. *Genome Announc.* 2(3):e00489-14. doi:10.1128/genomeA.00489-14.

Copyright © 2014 Adam et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to James T. Tambong, james.tambong@agr.gc.ca.

Pantoea ananatis is pathogenic on several economically important agricultural crops including pineapple (1), onions (2), maize (3), and rice (4). To date, five *P. ananatis* genomes have been published but none is the type strain. We report the draft genome sequence of the type strain of *P. ananatis* LMG 2665, the bacterial causal agent of the fruitlet rot of pineapple.

The draft genome of *P. ananatis* LMG 2665^T was determined by paired-end sequencing using an Illumina HiSeq 2500 with TrueSeq V3 chemistry at the National Research Council Canada (Saskatoon, Saskatchewan, Canada). A total of 18,809,898 pair-end reads, each 101 bp in length, totaling 1,899,799,698 bp, were obtained from 300-bp inserts. Quality checking using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) showed that the reads are of sufficiently good quality such that no further trimming or error correction was required. Initial *de novo* assembly, using ABYSS v1.3.6 (5), produced 88 contigs contained in 73 scaffolds, from which scaffolds with length <300 bp were removed. The remaining 29 scaffolds (minimum, 1,050 bp; maximum, 612,849 bp; N_{50} , 307,200 bp; total size, 4,980,044 bp; total number of unknown nucleotides [Ns], 201) were used for further analyses. SSPACE v2.0 (6) was applied on the resulting scaffolds to possibly extend and merge them into larger scaffolds based on read-pair information and short overlaps. This process reduced the number of scaffolds to 27 (minimum, 1050 bp; maximum, 612,849 bp; N_{50} , 307,200 bp; total size, 4,980,628 bp; total number of Ns, 203). GapFiller v1.11 (7) was then used to close the gaps between the short scaffolds, that are contained within the large 27 scaffolds, by replacing the unknown nucleotides (Ns) with true nucleotides based on read-pair information and short overlaps. The final draft genome consists of 27 scaffolds (minimum, 1,050 bp; maximum, 612,855 bp; N_{50} , 307,200 bp) totaling 4,980,528 bp with 50 Ns. The G+C content of the draft genome is 53.40% with an overall estimated coverage at 380×.

Mauve Contig Mover v2.3.1 (8) was applied to order the draft genome of the *P. ananatis* strain LMG 2665^T using *Pantoea ananatis* AJ13355 (accession no. NC_017531.1) as a reference genome. Automated annotation using the RAST annotation server (9) revealed that the draft genome of *P. ananatis* LMG 2665^T contains 4,787 predicted protein-coding sequences, of which 3,749 have assigned functions, 267 have proposed functions, and 771 have

been considered hypothetical proteins, respectively. The draft genome also contains 88 predicted noncoding RNAs including 62 tRNAs and 16 rRNAs. The number of gene copies encoding 16S rRNA, 5S rRNA, and 23S rRNA are 4, 7, and 5, respectively. Compared to RAST, Glimmer v3.02 (10) using open reading frames (ORFs) as a training set predicted 4,893 genes, whereas RNAmmer v1.2 (11) predicted 14 rRNAs containing 1 copy of 16S rRNA, 8 copies of 5S rRNA, and 5 copies of 23S rRNA genes. However, tRNAscan-SE v1.3.1 (12) predicted 72 tRNAs genes.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [JFZU00000000](https://www.ncbi.nlm.nih.gov/nuccore/JFZU00000000). The version described in this paper is the first version, JFZU01000000.

ACKNOWLEDGMENTS

This study was funded by the Canadian Safety and Security Program (CSSP) and Agriculture and Agri-Food Canada through the Agri-Food pathogen detection and Next-Gen sequencing project CRTI 09-462RD and Mycology and Bacteriology Biosystematics project no. 1800.

REFERENCES

- Serrano FB. 1928. Bacterial fruitlet brown rot on pineapple in the Philippines. *Philipp. J. Sci.* 36:271–305.
- Gitaitis RD, Gay JD. 1997. First report of a leaf blight, seed stalk rot, and bulb decay of onion by *Pantoea ananatis* in Georgia. *Plant Dis.* 81:1096. <http://dx.doi.org/10.1094/PDIS.1997.81.9.1096B>.
- Paccola-Meirelles LD, Ferreira AS, Meirelles WF, Marriel IE, Casela CR. 2001. Detection of a bacterium associated with a leaf spot disease of maize in Brazil. *J. Phytopathol.* 149:275–279. <http://dx.doi.org/10.1046/j.1439-0434.2001.00614.x>.
- Cother EJ, Reinke R, McKenzie C, Lanoiselet VM, Noble DH. 2004. An unusual stem necrosis of rice caused by *Pantoea ananatis* and the first record of this pathogen on rice in Australia. *Austral. Plants Pathol.* 33:495–503. <http://dx.doi.org/10.1071/AP04053>.
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. 2009. ABYSS: A parallel assembler for short read sequence data. *Genome Res.* 19:1117–1123. <http://dx.doi.org/10.1101/gr.089532.108>.
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
- Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.

8. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics* 25:2071–2073. <http://dx.doi.org/10.1093/bioinformatics/btp356>.
9. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
10. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 27: 4636–4641. <http://dx.doi.org/10.1093/nar/27.23.4636>.
11. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
12. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964. <http://dx.doi.org/10.1093/nar/25.5.0955>.